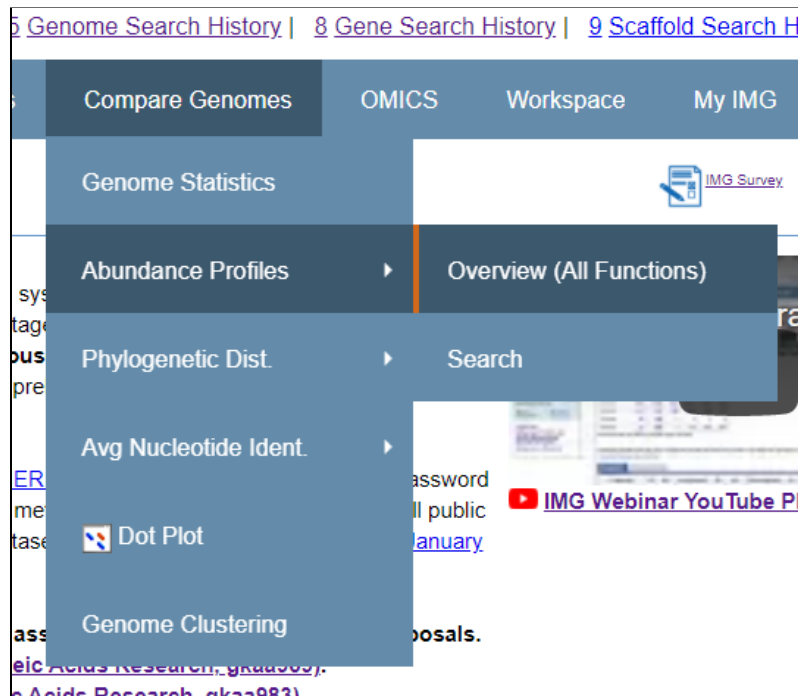


Abundance Profile Overview Tool

Purpose. Examine relative abundance of all protein families (COGs and Pfams) and functional families (Enzymes) in isolate genomes, metagenomes and metagenome bins (or workspace scaffold sets).

Menu Navigation: Compare Genomes → Abundance Profiles → Abundance Profile Overview



Functionality. Select the type of format for displaying the results (“Heat Map” or “Matrix”), the type of protein/functional families (COG, Pfam, Enzyme), normalization method, and a set of isolate genomes, metagenomes and/or metagenome bins (or workspace scaffold sets).

Abundance Profile Overview Form

Display options

Heat Map: the abundance of protein/functional families will be displayed as a heat map with red corresponding to the most abundant families.

Matrix: the abundance of protein/functional families is displayed in a tabular format.

Abundance Profile Overview

Limitations:

- Please limit a scaffold set size (less than 20) for real time computation
- Please try to use Submit as Computation Job in [\(IMG/MER\)](#) when using scaffold sets

Display Options:

— Output display option

Output Type

☐ Heat Map

OR

☒ Matrix

Normalization Method ³

☐ None

☐ Scale for genome size

☒ Gene count

☐ Include all rows, including those without hits

☐ Estimated gene copies ²

Enter matching text for highlighting clusters/rows (E.g., "kinase")

Function:

- ☒ COG
- ☐ Enzyme
- ☐ KO
- ☐ Pfam
- ☐ TIGRfam

Sequencing Status

All Finished, Permanent Draft and Draft

Domain

Genome Cart

☒ List
☐ Tree

Show

Selected: 1

Search for: <enter a genome name to search>

Butyrivibrio fibrisolvens 16/4 (B) [F]
Butyrivibrio fibrisolvens DSM 3071 (B) [P]
Butyrivibrio fibrisolvens YRB2005 (B) [P]
Butyrivibrio proteoclasticus B316 (B) [F]
Escherichia coli O44:H18 042 (EAEC) (B) [F]
Freshwater sediment methanotrophic microbial communities from Lake

Add Genomes

Add >
Add All >>

< Remove
<< Remove All

Selected Genomes

Please select 1 to 100 genomes. 2 selected

Butyrivibrio fibrisolvens 16/4 (B) [F]
Freshwater sediment methanotrophic microbial comm
Butyrivibrio
Abundance_test_merfs

Upload Sets

MERFS Metagenome: Assembled

Go

Real time computation

Submit as Computation Job

Submit a computation to run in the background.
☒ Save as a new job with name: abundance test2
☐ Replace the selected job:

Comment:

Submit Computation

Submit for background computation (recommended)

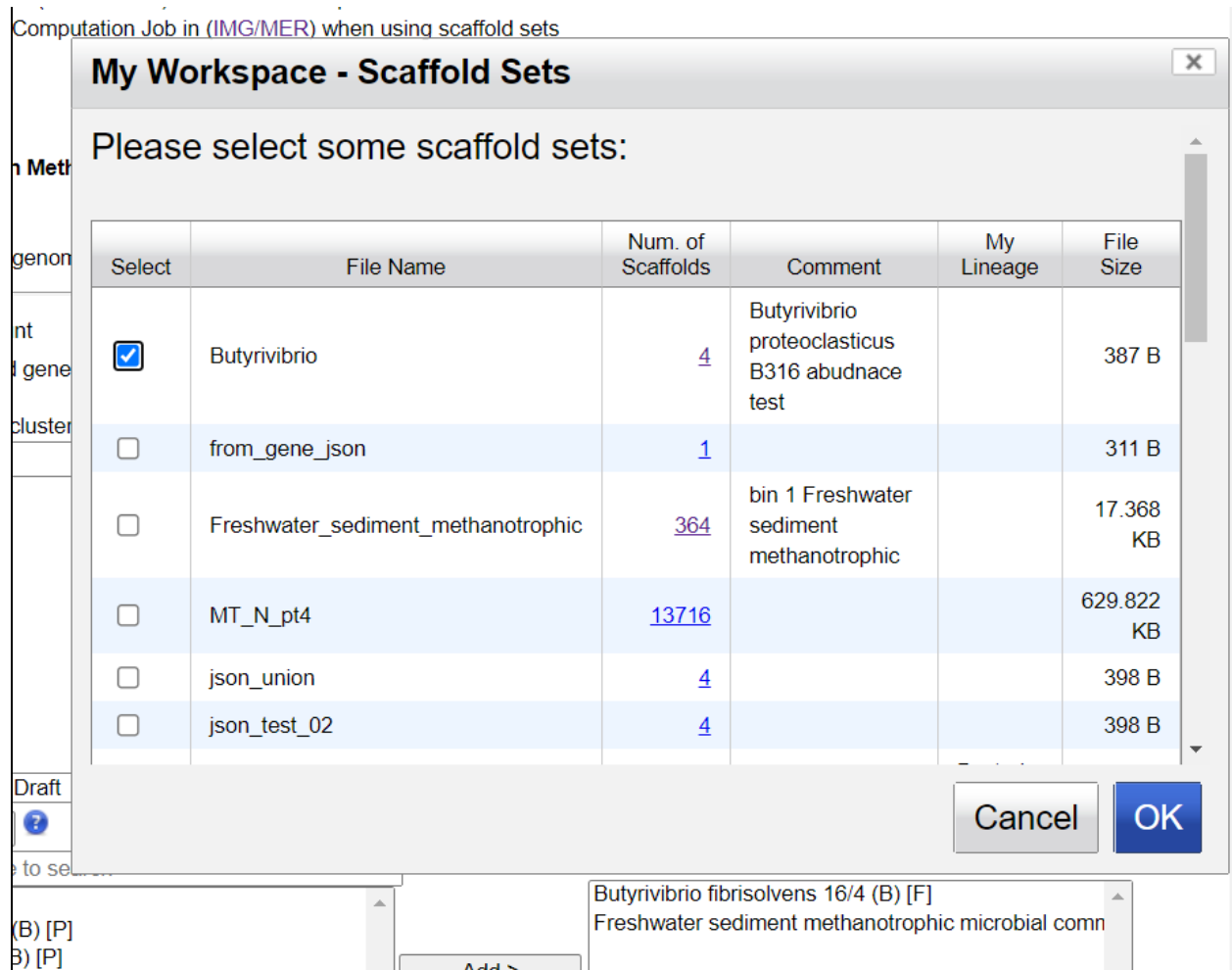
Data Set selection

Selecting Genomes:

Click on one or more genome names, and then click the “Add” button to confirm selection.

NEW Scaffold Set Selector:

Click the “Upload Sets” to view a select scaffold sets dialog.



Real time vs Background Computation

Pressing the “Go” button runs the computation in real time with some limitations:

1. Selecting a maximum of 100 genomes.
2. Scaffold set size should have less than 20 scaffolds.

Selecting too many data sets may result in timeouts.

We recommend “**Background Computation**” which is only available in IMG/MER.

Using background computation you can select more data sets and larger scaffold sets. Your results will be saved in your workspace.

Output Results

If the “**Matrix**” output is selected, the abundance of protein/functional families is displayed in a tabular format, with each row corresponding to a family and each cell containing the number of genes associated with a family for a specific genome or

metagenome. Click on the cell in order to retrieve the list of genes assigned to this particular family in a genome or metagenome. Families of interest can be selected for inclusion into the Function Cart.

The results in “Matrix” format can be exported to a tab delimited Excel file.

Abundance Profile Overview Results (Gene Count)

1 - [Butyrivibrio fibrisolvens 16/4](#)
 2 - [Freshwater sediment methanotrophic microbial communities from Lake Washington under simulated oxygen tension - Sediment Metagenome 20_HOW5 \(SPAdes\)](#)
 3 - [Butyrivibrio](#)
 4 - [Abundance_test_merfs](#)

hint: - Click on count to see constituent genes.
 - Counts for cell coloring: white < 1, bisque = 1-5, pink = 6-10, yellow > 10..

[Add Selected to Function Cart](#) [Select All](#) [Clear All](#)

Filter column: Filter [Apply](#) [?](#)

[Export](#) Page 1 of 35 << first < prev 1 2 3 4 5 6 7 8 9 10 next > last >> 100 ▾

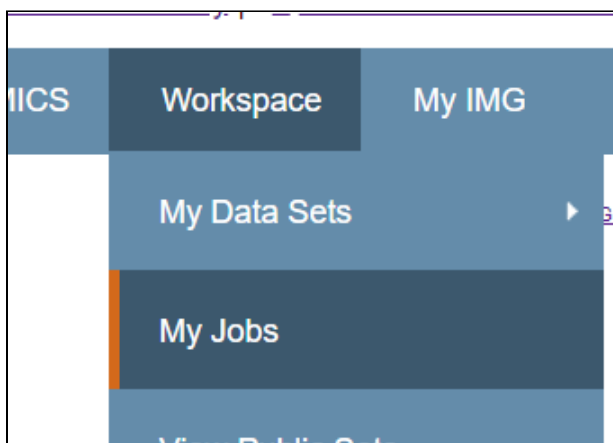
[Column Selector](#) [Select Page](#) [Deselect Page](#)

| Select | Function Id ▲ | Function Name | Butyrivibr fibrisolve 16 650377919 | Freshwater sediment methanotro 3300024970 | Butyrivibrio | Abundance_test_merfs |
|--------------------------|---------------|---|---|--|--------------|----------------------|
| <input type="checkbox"/> | COG0001 | Glutamate-1-semialdehyde aminotransferase | 0 | 59 | 0 | 1 |
| <input type="checkbox"/> | COG0002 | N-acetyl-gamma-glutamylphosphate reductase | 1 | 58 | 1 | 0 |
| <input type="checkbox"/> | COG0003 | Anion-transporting ATPase, ArsA/GET3 family | 0 | 21 | 0 | 0 |

Workspace Job Output example

For background computation you will receive an email when your results have been completed. All results will be saved in your workspace.

Workspace Menu navigation:



Job summary list:

List of all my jobs.

| Program | Description | Count |
|----------------------------|-------------|-------|
| Abundance Profile Overview | | 10 |

[View list of jobs](#)

Abundance Profile Overview list:

Clicking the “10” from above, you will get all of your “Abundance Profile Search” jobs.

Click the “completed” url to view your results.

| | | | | | | | |
|--------------------------|-----------------|----------------------------|---------------------|-----------------|------------|---------------------|---------------------------|
| <input type="checkbox"/> | abundance_test2 | Abundance Profile Overview | 2021/04/13 11:49:30 | overview search | 479.503 KB | 2021/04/19 21:10:02 | completed |
|--------------------------|-----------------|----------------------------|---------------------|-----------------|------------|---------------------|---------------------------|

[View Results](#)

Results page:

Abundance Profile Overview Results

Job: [abundance_test2](#) — Job results
Comment: overview search

1 - [Butyrivibrio fibrisolvens 16/4](#)
2 - [Freshwater sediment methanotrophic microbial communities from Lake Washington under simulated oxygen tension - Sediment Metagenome 20_HOW5 \(SPAdes\)](#)
3 - [Butyrivibrio](#)
4 - [Abundance_test_merfs](#)

[Add Selected to Function Cart](#) [Select All](#) [Clear All](#)

Filter column: Function Id Filter text [Apply](#)

[Export](#) Page 1 of 35 << first < prev 1 2 3 4 5 6 7 8 9 10 next > last >> 100

[Column Selector](#) [Select Page](#) [Deselect Page](#)

| Select | Function Id | Function Name | Butyrivibrio fibrisolve 16 650377919 | Freshwater sediment methanotro 3300024970 | Butyrivibrio | Abundance_test_merfs |
|--------------------------|-------------|--|--------------------------------------|---|--------------|----------------------|
| <input type="checkbox"/> | COG0001 | Glutamate-1-semialdehyde aminotransferase | 0 | 59 | 0 | 1 |
| <input type="checkbox"/> | COG0002 | N-acetyl-gamma-glutamylphosphate reductase | 1 | 58 | 1 | 0 |

Heat Map Output

For “**Heat Map**” output, the abundance of protein/functional families will be displayed as a heat map with red corresponding to the most abundant families. Each column on the map corresponds to a genome, metagenome or scaffold set, each row – to a family; mouse over each cell to see the count of a particular family in a genome/metagenome. Clicking on the ID of the family displayed right to the column will add the corresponding family to the Function Cart; clicking on the cell will retrieve the list of genes assigned to this particular family in this genome or metagenome. The map can be sorted according to the abundance by clicking on the corresponding column header.

Abundance Profile

Mouse over labels to see additional information.

Clicking on the column number will sort rows for that column in

Clicking on row cluster ID will add the cluster to the appropriate

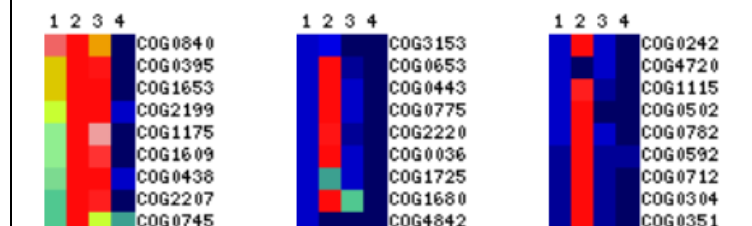
Mouse over heat map to see gene counts. Clicking in the heat

1 - [Butyrivibrio fibrisolvens 16/4](#)

2 - [Freshwater sediment methanotrophic microbial community](#)

3 - [Butyrivibrio](#)

4 - [Abundance_test_merfs](#)



Notes. This analysis does not include the read depth coverage of each gene when counting family abundance. Beware when comparing high-complexity metagenomes with very low degree of assembly (e.g., soil) with low-complexity well-assembled metagenomes, such as AMD sample, since each gene in the latter may correspond to many reads.